Smoothness Characterization and Stability of Nonlinear and Non-Separable Multi-scale Representations

Basarab Matei, Sylvain Meignen, Anastasia Zakharova

Smoothness Characterization and Stability of Nonlinear and Non-Separable Multi-Scale Representations

Basarab Matei[a], Sylvain Meignen*[b], Anastasia Zakharova[c]

[a] *LAGA Laboratory, Paris XIII University, France,*
Tel:0033-1-49-40-35-71
FAX:0033-4-48-26-35-68
E-mail: matei@math.univ-paris13.fr

[b] *LJK Laboratory, University of Grenoble, France*
Tel:0033-4-76-51-43-95
FAX:0033-4-76-63-12-63
E-mail: sylvain.meignen@imag.fr

[c] *LJK Laboratory, University of Grenoble, France*
Tel:0033-4-76-51-43-95
FAX:0033-4-76-63-12-63
E-mail: anastasia.zakharova@imag.fr

Preprint submitted to Journal of Approximation Theory *December 21, 2010*

Abstract

The aim of the paper is the construction and the analysis of nonlinear and non-separable multi-scale representations for multivariate functions. The proposed multi-scale representation is associated with a non-diagonal dilation matrix M. We show that the smoothness of a function can be characterized by the rate of decay of its multi-scale coefficients. We also study the stability of these representations, a key issue in the designing of adaptive algorithms.
Keywords: Nonlinear Multiscale approximation, Besov Spaces, Stability.
2000 MSC: 41A46, MSC 41A60, MSC 41A63

1. Introduction

A multi-scale representation of an abstract object v (e.g. a function representing the grey level of an image) is defined as $\mathcal{M}v := (v^0, d^0, d^1, d^2, \cdots)$, where v^0 is the coarsest approximation of v in some sense and d^j, with $j \geq 0$, are additional detail coefficients representing the fluctuations between two successive levels. Several strategies exist to build such representations: wavelet basis, lifting schemes and also the discrete framework of Harten [9]. Using a wavelet basis, we compute $(v^0, d^0, d^1, d^2, \cdots)$ through linear filtering and thus the multi-scale representation corresponds to a change of basis. Although wavelet bases are optimal for one-dimensional functions, this is no longer the case for multivariate objects such as images where the presence of singularities requires special treatments. The approximation property of wavelet bases and their use in image processing are now well understood (see [6] and [15] for details).

Overcoming this "curse of dimensionality" for wavelet basis was in the past decade the subject of active research. We mention here several strategies developed from the wavelets theory: the curvelets transforms [3], the directionlets transforms [7] and the bandelets transform [13]. Another approach proposed in [16] and studied in [2] uses the discrete framework of Harten, which allows a better treatment of singularities and consequently better approximation results.

The applications of all these methods to image processing are numerous: let us mention some of these works in [2], [1] and [4]. In [2], the extension of univariate methods using tensor product representations is studied. Although this extension is natural and simple, the results are not optimal.

We propose in the present paper a new nonlinear multi-scale representation based on the general framework of A. Harten (see [9] and [11]). Our representation is non-separable and is associated with a non-diagonal dilation matrix M. The use of non-diagonal dilation matrices is motivated by better image compression performances ([5] and [14]). Since the details are computed adaptively, the multi-scale representations is completely nonlinear and is no more equivalent to a change of basis. To study these representations, we develop some new analysis tools. In particular we make extensive use of mixed finite differences and of associated joint spectral radii to generalize the existing convergence and stability results based on a tensor product approach. The smoothness characterization is based both on direct and inverse theorems. To prove the direct theorem we use the polynomial reproduction and we assume that the dilation matrix is isotropic. For the inverse theorem the assumption on the isotropy of the matrix is not necessary. We prove that our representations give the same approximation order as for wavelet basis. This strategy is fruitful in applications since it allows to cope up with the deficiencies of wavelet bases without loosing the approximation order. More precisely, in this paper we show that the convergence and the stability in L^p and Besov spaces of our representations can be obtained under the same hypothesis on the joint spectral radii associated to mixed finite differences. This was not the case in previous one-dimensional or tensor product studies [20] and [17], where the joint spectral associated to finite differences had to be lower than one to ensure stability. The outline of the paper is the following. After having introduced nonlinear and non-separable multi-scale representations, we give an illustration on image compression of the improvement brought about the use of non-diagonal matrix in the multi-scale representation. Extending the results of [17], we characterize the smoothness of a function v belonging to some Besov spaces by means of the decay of the detail coefficients of its nonlinear and non-separable multiscale representation (section 4 and 5). We finally study the stability of this multi-scale representation in section 6 (for similar, one-dimensional results see [20] and [17]).

2. Multi-scale Representations on $\mathbb{R}^d$

For the reader convenience, we recall the construction of linear multi-scale representations based on multiresolution analysis (MRA). To this end, let M be a $d \times d$ dilation matrix.

Definition 1. *A multiresolution analysis of V is a sequence $(V_j)_{j\in\mathbb{Z}}$ of closed subspaces of V satisfying the following properties:*

1. *The subspaces are embedded: $V_j \subset V_{j+1}$;*
2. *$f \in V_j$ if and only if $f(M.) \in V_{j+1}$;*
3. *$\overline{\cup_{j\in\mathbb{Z}} V_j} = V$;*
4. *$\cap_{j\in\mathbb{Z}} V_j = \{0\}$;*
5. *There exists a compactly supported function $\varphi \in V_0$ such that the family $\{\varphi(\cdot - k)\}_{k\in\mathbb{Z}^d}$ forms a Riesz basis of V_0.*

The function φ is called the *scaling function*. Since $V_0 \subset V_1$, φ satisfies the following equation:

$$\varphi = \sum_{k\in\mathbb{Z}^d} g_k \varphi(M \cdot -k), \text{ with } \sum_k g_k = m := |\det(M)|. \tag{1}$$

To get the approximation of a given function v at level j, we consider a compactly supported function $\tilde{\varphi}$ *dual* to φ (i.e. for all $k, n \in \mathbb{Z}^d$ $< \tilde{\varphi}(\cdot - n), \varphi(\cdot - k) >= \delta_{n,k}$, where $\delta_{n,k}$ denotes the Kronecker symbol and $< \cdot, \cdot >$ the inner product on V), which also satisfies a scaling equation

$$\tilde{\varphi} = \sum_{n \in \mathbb{Z}^d : \|n\|_\infty \leq P} \tilde{h}_n \tilde{\varphi}(M \cdot - n), \text{ with } \sum_k \tilde{h}_k = m. \tag{2}$$

The approximation v_j of v we consider is then obtained by projection of v on V_j as follows:

$$v_j = \sum_{n \in \mathbb{Z}^d} v_n^j \varphi(M^j \cdot - n). \tag{3}$$

where

$$v_n^j = \int v(x) m^j \tilde{\varphi}(M^j x - n) dx, \quad n \in \mathbb{Z}^d. \tag{4}$$

Multi-scale representations based on specific choice for $\tilde{\varphi}$ are commonly used in image processing and numerical analysis. We mention two of them: the first one is the point-values case obtained when $\tilde{\varphi}$ is the Dirac distribution and the second one is the cell average case obtained when $\tilde{\varphi}$ is the indicator function of some domain on $\mathbb{R}^d$. In the theoretical study that follows, we assume that the data are obtained through a projection of a functional v as in (4).

A strategy which allows to build nonlinear multi-scale representations based on such a projection can be done in terms of a very general discrete framework using the concept of inter-scale operators introduced by A. Harten in [9], which we now recall. Assume that we have two inter-scale discrete operators associated to this sequence: the *projection* operator P_{j-1}^j and the *prediction* operator P_j^{j-1}. The projection operator P_{j-1}^j acts from fine to coarse levels, that is, $v^{j-1} = P_{j-1}^j v^j$. This operator is assumed to be *linear*.

The *prediction* operator P_j^{j-1} acts from coarse to fine levels. It computes the 'approximation' $\hat{v}^j$ of v^j from the vector $(v_k^{j-1})_{k\in\mathbb{Z}^d}$ which is associated to $v_{j-1} \in V_{j-1}$:

$$\hat{v}^j = P_j^{j-1} v^{j-1}.$$

This operator may be *nonlinear*. Besides, we assume that these operators satisfy the *consistency* property:

$$P_{j-1}^j P_j^{j-1} = I, \tag{5}$$

i.e., the projection of $\hat{v}^j$ coincides with v^{j-1}. Having defined the prediction error $e^j := v^j - \hat{v}^j$, we obtain a redundant representation of vector v^j:

$$v^j = \hat{v}^j + e^j. \tag{6}$$

By the consistency property, one has

$$P_{j-1}^j e^j = P_{j-1}^j v^j - P_{j-1}^j \hat{v}^j = v^{j-1} - v^{j-1} = 0.$$

Hence, $e^j \in \mathrm{K}er(P_{j-1}^j)$. Using a basis of this kernel , we write the error e^j in a non-redundant way and get the detail vector d^{j-1}. The data v^j is thus completely equivalent to the data (v^{j-1}, d^{j-1}). Iterating this process from the initial data v^J, we obtain its *nonlinear multi-scale representation*

$$\mathcal{M}v^J = (v^0, d^0, \ldots, d^{J-1}). \tag{7}$$

From here on, we assume the equivalence

$$\|e^j\|_{\ell^p(\mathbb{Z}^d)} \sim \|d^{j-1}\|_{\ell^p(\mathbb{Z}^d)}. \tag{8}$$

Since the details are computed adaptively, the underlying multi-scale representation is nonlinear and no more equivalent to a change of basis. Moreover, the discrete setting used here is not based on the study of scaling equations as for wavelet basis, which implies that the results of wavelet theory cannot be used directly in our analysis. Note also that the projection operator is completely characterized by the function $\tilde{\varphi}$. Namely, if we consider the discretization defined by (4) then, in view of (2), we may write the projection operator as follows:

$$v_k^{j-1} = m^{-1} \sum_{\|n\|_\infty \leq P} \tilde{h}_n v^j_{Mk+n} = m^{-1} \sum_{\|n-Mk\|_\infty \leq P} \tilde{h}_{n-Mk} v^j_n := (P^j_{j-1} v^j)_k. \quad (9)$$

To describe the prediction operator, for every $w \in \ell^\infty(\mathbb{Z}^d)$ we consider a linear operator $S(w)$ defined on $\ell^\infty(\mathbb{Z}^d)$ by

$$(S(w)u)_k := \sum_{l \in \mathbb{Z}^d} a_{k-Ml}(w) u_l, \quad k \in \mathbb{Z}^d. \quad (10)$$

Note that the coefficients $a_k(w)$ depend on w. We assume that S is local:

$$\exists K > 0 \quad \text{such that} \quad a_{k-Ml}(w) = 0 \text{ if } \|k - Ml\|_\infty > K \quad (11)$$

and that $a_k(w)$ is bounded independently of w:

$$\exists C > 0 \quad \text{such that} \quad \forall w \in \ell^\infty(\mathbb{Z}^d) \quad \forall k, l \in \mathbb{Z}^d \quad |a_{k-Ml}(w)| < C. \quad (12)$$

Remark 2.1. *From* (12) *it immediately follows that for any* $p \geq 1$ *the norms* $\|S(w)\|_{\ell^p(\mathbb{Z}^d)}$ *are bounded independently of* w*.*

The *quasi-linear* prediction operator is then defined by

$$\hat{v}^j = P_j^{j-1} v^{j-1} = S(v^{j-1}) v^{j-1}. \quad (13)$$

If for all $k, l \in \mathbb{Z}^d$ and all $w \in \ell^\infty(\mathbb{Z}^d)$ we put $a_{k-Ml}(w) = g_{k-Ml}$, where g_{k-Ml} is defined by the scaling equation (1), we get the so-called *linear* prediction operator. In the general case, the prediction operator P_j^{j-1} can be viewed as a perturbation of the *linear* prediction operator due to the *consistency* property; that is why we will call it a *quasi-linear* prediction operator. The operator-valued function which associates to any w an operator $S(w)$ is called a *quasi-linear* prediction rule.

For what follows, we need to introduce the notion of polynomial reproduction for *quasi-linear* prediction rules. A polynomial q of degree N is defined as a linear combination $q(x) = \sum_{|n|\leq N} c_n x^n$. Let us denote by Π the linear space of all polynomials, by Π_N the linear space of all polynomials of degree N. With this in mind, we have the following definition for polynomial reproduction:

Definition 2.1. *We will say that the quasi-linear prediction rule $S(w)$ reproduces polynomials of degree N if for any $w \in \ell^\infty(\mathbb{Z}^d)$ and any $u \in \ell^\infty(\mathbb{Z}^d)$ such that $u_k = p(k)$ $\forall k \in \mathbb{Z}^d$ for some $p \in \Pi_N$, we have:*

$$(S(w)u)_k = p(M^{-1}k) + q(k),$$

where $\deg(q) < \deg(p)$*. If $q = 0$, we say that the quasi-linear prediction rule S exactly reproduces polynomials of degree N.*

Note that the property is required *for any data* w, and not only for $w = u$. In the following, we will consider dilation matrices to define inter-scale operators. A dilation matrix is an invertible integer-valued matrix M satisfying $\lim\limits_{n\to\infty} M^{-n} = 0$, and $m := |\det(M)|$.

3. Improvement Brought about Non-separable Representations for Image Compression

In this section , we illustrate on an example the potential interest of non-separable representations for image compression. The interested reader may consult [5] and [14] for further details. The dilation matrix we use here is the quincunx matrix defined by:

$$M = \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix},$$

whose coset vectors are $\varepsilon_0 = (0,0)^T$ and $\varepsilon_1 = (0,1)^T$. We consider an interpolatory multi-scale representation which implies that $v_k^j = v(M^{-j}k)$ (i.e. $\tilde{\varphi}$ is the Dirac function). By construction $v_{Mk}^j = v_k^{j-1}$, so we only need to predict $v(M^{-j}k + \varepsilon_1)$. To do so, we define four polynomials of degree 1 (i.e. $a + bx + cy$) interpolating v on the following stencils:

$$\begin{aligned} V_k^{j,1} &= M^{-j+1}\{k, k+e_1, k+e_2\} \\ V_k^{j,2} &= M^{-j+1}\{k, k+e_1, k+e_1+e_2\} \\ V_k^{j,3} &= M^{-j+1}\{k+e_1, k+e_2, k+e_1+e_2\} \\ V_k^{j,4} &= M^{-j+1}\{k, k+e_2, k+e_1+e_2\}, \end{aligned}$$

which in turn entails the following two predictions for $v(M^{-j}k + \varepsilon_1)$:

$$\hat{v}_{Mk+\varepsilon_1}^{j,1} = \tfrac{1}{2}(v_k^{j-1} + v_{k+e_1+e_2}^{j-1}) \tag{14}$$

$$\hat{v}_{Mk+\varepsilon_1}^{j,2} = \tfrac{1}{2}(v_{k+e_1}^{j-1} + v_{k+e_2}^{j-1}). \tag{15}$$

We now show that to choose between the two predictions (14) and (15) appropriately improves the compression performance on natural images. The cost function we use to make the choice of stencil is as follows:

$$C^j(k) = \min(|v^{j-1}_{k+e_1} - v^{j-1}_{k+e_2}|, |v^{j-1}_{k} - v^{j-1}_{k+e_1+e_2}|).$$

When the minimum of $C^j(k)$ corresponds to the first (resp. second) argument, the prediction (15) (resp. (14)) is used. The motivation for the choice of such a cost function is the following: when an edge intersect the cell Q^{j-1}_k defined by the vertices $M^{-j+1}\{k, k+e_1, k+e_2, k+e_1+e_2\}$, several cases may happen:

1. either the edge intersects $[M^{-j+1}k, M^{-j+1}(k+e_1+e_2)]$ and $[M^{-j+1}(k+e_1), M^{-j+1}(k+e_2)]$ in which case no direction is favored.
2. or the edge intersects $[M^{-j+1}k, M^{-j+1}(k+e_1+e_2)]$ or $[M^{-j+1}(k+e_1), M^{-j+1}(k+e_2)]$, in which case the prediction operator favors the direction which is not intersected by the edge (this will lead to a better prediction).

When Q^{j-1}_k is not intersected by an edge, the gain between choosing one direction or the other is negligible and, in that case, we will apply predictions (14) and (15) successively. It thus remains to determine when a cell is intersected by an edge. An edge-cell is determined by the following condition:

$$\begin{aligned} \underset{k'=k,k+e_1+e_2,k-e_1-e_2}{\operatorname{argmin}} |v^{j-1}_{k'} - v^{j-1}_{k'+e_1+e_2}| = 1 \text{ or} \\ \underset{k'=k,k+e_1-e_2,k-e_1+e_2}{\operatorname{argmin}} |v^{j-1}_{k'+e_1} - v^{j-1}_{k'+e_2}| = 1 \qquad (C), \end{aligned}$$

which means the first order differences are locally maximum in the direction of prediction. Then, to encode the representation, we use an adapted version to our context of the EZW (Embedded Zero-tree Wavelet) [21]. To simplify, consider a $N \times N$ image with $N = 2^J$ and J even, then d^j (defined in (7)) is

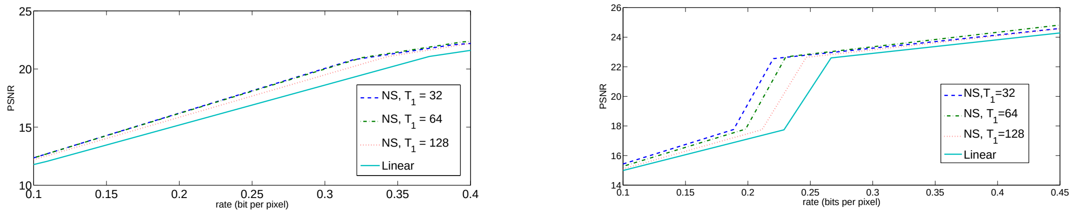
(A)

(B)

Figure 1: (A): a 256×256 Lena image, (B): a 256×256 peppers image

associated to a $2^j \times 2^{j+1}$ matrix of coefficients when j is odd and to a $2^j \times 2^j$ matrix of coefficients when j is even. We denote by T_1^j the number of lines of the matrix associated to d^j. We display the compression results for the 256×256 images of Figure 1 on Figure 2 (A) and (B). We apply nonlinear prediction only on edge-cells detected using (C) and only for the subspaces V_j such that $T_1^j \geq T_1$. It means that for instance, if we take $T_1 = 64$ and $N = 256$, we only predict nonlinearly the last finest four detail coefficients subspaces. A typical compression result is labeled by NS where NS stands for non-separable.

In any tested cases (i.e. for several T_1), using nonlinear predictions leads better compression results than the one based on a diagonal dilation matrix. These encouraging results motivate a deeper study of nonlinear and non-separable multi-scale representations which we now carry out.



(A) (B)

Figure 2: (A): linear prediction (solid line) and NS (non-separable) prediction on edge-cells computed using (C) and for varying T_1 for the image of Lena, (B):idem but for the image of peppers

4. Notations and Generalities

We start by introducing some notations that will be used throughout the paper. Let us consider a multi-index $\mu = (\mu_1, \mu_2, \ldots, \mu_d) \in \mathbb{N}^d$ and a vector $x = (x_1, x_2, \ldots, x_d) \in \mathbb{R}^d$. We define $|\mu| = \sum_{i=1}^{d} \mu_i$ and $x^\mu = \prod_{i=1}^{d} {x_i}^{\mu_i}$. For two multi-indices $m, \mu \in \mathbb{N}^d$ we define

$$\begin{pmatrix} \mu \\ m \end{pmatrix} = \begin{pmatrix} \mu_1 \\ m_1 \end{pmatrix} \cdots \begin{pmatrix} \mu_d \\ m_d \end{pmatrix}.$$

For a fixed integer $N \in \mathbb{N}$, we define

$$q_N = \#\{\mu, |\mu| = N\} \tag{16}$$

where $\#Q$ stands for the cardinal of the set Q. The space of bounded sequences is denoted by $\ell^\infty(\mathbb{Z}^d)$ and $\|u\|_{\ell^\infty(\mathbb{Z}^d)}$ is the supremum of $\{|u_k| : k \in \mathbb{Z}^d\}$. As usual, let $\ell^p(\mathbb{Z}^d)$ be the Banach space of sequences u on $\mathbb{Z}^d$ such that $\|u\|_{\ell^p(\mathbb{Z}^d)} < \infty$, where

$$\|u\|_{\ell^p(\mathbb{Z}^d)} := \left(\sum_{k \in \mathbb{Z}^d} |u_k|^p \right)^{\frac{1}{p}} \text{ for } 1 \leq p < \infty.$$

We denote by $L^p(\mathbb{R}^d)$,the space of all measurable functions v such that $\|v\|_{L^p(\mathbb{R}^d)} < \infty$, where

$$\|v\|_{L^p(\mathbb{R}^d)} := \left(\int_{\mathbb{R}^d} |v(x)|^p dx\right)^{\frac{1}{p}} \text{ for } 1 \leq p < \infty,$$
$$\|v\|_{L^\infty(\mathbb{R}^d)} := \operatorname*{ess\,sup}_{x\in\mathbb{R}^d} |v(x)|.$$

Throughout the paper, the symbol $\|\cdot\|_\infty$ is the sup norm in $\mathbb{Z}^d$ when applied either to a vector or a matrix. Let us recall that, for a function v, the finite difference of order $N \in \mathbb{N}$, in the direction $h \in \mathbb{R}^d$ is defined by:

$$\nabla_h^N v(x) := \sum_{k=0}^{N} (-1)^k \begin{pmatrix} N \\ k \end{pmatrix} v(x+kh).$$

and the mixed finite difference of order $n = (n_1, \dots, n_d) \in \mathbb{N}^d$ in the direction $h = (h_1, \dots, h_d) \in \mathbb{R}^d$ by:

$$\nabla_h^n v(x) := \nabla_{h_1e_1}^{n_1} \cdot \ldots \cdot \nabla_{h_de_d}^{n_d} v(x) = \sum_{k_1,\dots,k_d=0}^{\max(n_1,\dots,n_d)} (-1)^{|n|} \begin{pmatrix} n \\ k \end{pmatrix} v(x+k\cdot h),$$

where $k \cdot h := \sum_{i=1}^{d} k_ih_i$ is the usual inner product while $(e_1, \dots, e_d)$ is the canonical basis on $\mathbb{Z}^d$. For any invertible matrix B we put

$$\nabla_B^n v(x) := \nabla_{Be_1}^{n_1} \cdot \ldots \cdot \nabla_{Be_d}^{n_d} v(x).$$

Similarly, we define $D^\mu v(x) = D_1^{\mu_1} \cdots D_d^{\mu_d} v(x)$, where D_j is the differential operator with respect to the jth coordinate of the canonical basis. For a sequence $(u_p)_{p\in\mathbb{Z}^d}$ and a multi-index n, we will use the mixed finite differences of order n defined by the formulae

$$\nabla^n u : = \nabla_{e_1}^{n_1} \nabla_{e_2}^{n_2} \cdot \ldots \cdot \nabla_{e_d}^{n_d} u,$$

where $\nabla_{e_i}^{n_i}$ is defined recursively by

$$\nabla_{e_i}^{n_i} u_k = \nabla_{e_i}^{n_i-1} u_{k+e_i} - \nabla_{e_i}^{n_i-1} u_k.$$

Then, we put for $n \in \mathbb{N}$:

$$\Delta^N u : \ = \{\nabla^n u, |n| = N, \ n \in \mathbb{N}^d\}.$$

We end this section with the following remark on notations: for two positive quantities A and B depending on a set of parameters, the relation $A \lesssim B$ implies the existence of a positive constant C, independent of the parameters, such that $A \leq CB$. Also $A \sim B$ means $A \lesssim B$ and $B \lesssim A$.

4.1. Besov Spaces

Let us recall the definition of Besov spaces. Let $p, q \geq 1$, s be a positive real number and N be any integer such that $N > s$. The Besov space $B^s_{p,q}(\mathbb{R}^d)$ consists of those functions $v \in L^p(\mathbb{R}^d)$ satisfying

$$(2^{js}\omega_N(v, 2^{-j})_{L^p})_{j\geq 0} \in \ell^q(\mathbb{Z}^d),$$

where $\omega_N(v,t)_{L^p}$ is the modulus of smoothness of v of order $N \in \mathbb{N} \setminus \{0\}$ in $L^p(\mathbb{R}^d)$:

$$\omega_N(v,t)_{L^p} = \sup_{\substack{h\in\mathbb{R}^d \\ \|h\|_2 \leq t}} \|\nabla_h^N v\|_{L^p(\mathbb{R}^d)}, \quad t \geq 0,$$

where $\|.\|_2$ is the Euclidean norm. The norm in $B^s_{p,q}(\mathbb{R}^d)$ is then given by

$$\|v\|_{B^s_{p,q}(\mathbb{R}^d)} := \|v\|_{L^p(\mathbb{R}^d)} + \|(2^{js}\omega_N(v, 2^{-j})_{L^p})_{j\geq 0}\|_{\ell^q(\mathbb{Z}^d)}.$$

Let us now introduce a new modulus of smoothness $\tilde{\omega}_N$ that uses mixed finite differences of order N:

$$\tilde{\omega}_N(v,t)_{L^p} = \sup_{\substack{n\in\mathbb{N}^d \\ |n|=N}} \sup_{\substack{h\in\mathbb{R}^d \\ \|h\|_2 \leq t}} \|\nabla_h^n v\|_{L^p(\mathbb{R}^d)}, \ \ t > 0.$$

It is easy to see that for any v in $L^p(\mathbb{R}^d)$, $\|\nabla_h^N v\|_{L^p(\mathbb{R}^d)} \lesssim \sum_{|n|=N} \|\nabla_h^n v\|_{L^p(\mathbb{R}^d)}$, thus $\omega_N(v,t)_{L^p} \lesssim \tilde{\omega}_N(v,t)_{L^p}$. The inverse inequality $\tilde{\omega}_N(v,t)_{L^p} \lesssim \omega_N(v,t)_{L^p}$ immediately follows from Lemma 4 of [20]. It implies that:

$$\|v\|_{B^s_{p,q}(\mathbb{R}^d)} \sim \|v\|_{L^p} + \|(2^{js}\tilde{\omega}_N(v,2^{-j})_{L^p})_{j\geq 0}\|_{\ell^q(\mathbb{Z}^d)}.$$

Going further, there exists a family of equivalent norms on $B^s_{p,q}(\mathbb{R}^d)$.

Lemma 4.1. *For all* $\sigma > 1$, $\|v\|_{B^s_{p,q}(\mathbb{R}^d)} \sim \|v\|_{L^p(\mathbb{R}^d)} + \|(\sigma^{js}\tilde{\omega}_N(v,\sigma^{-j})_{L^p})_{j\geq 0}\|_{\ell^q(\mathbb{Z}^d)}$.

Proof: Since $\sigma > 1$, for any $j > 0$ there exists $j' > 0$ such that $2^{j'} \leq \sigma^j \leq 2^{j'+1}$. According to this, we have the inequalities

$$2^{j's}\tilde{\omega}_N(v,2^{-j'-1})_{L^p} \leq \sigma^{js}\tilde{\omega}_N(v,\sigma^{-j})_{L^p} \leq 2^{(j'+1)s}\tilde{\omega}_N(v,2^{-j'})_{L^p},$$

from which the norm equivalence follows. □

5. Smoothness of Nonlinear Multi-scale Representations

In this section, we prove the equivalence between the norm of a function v belonging to $B^s_{p,q}(\mathbb{R}^d)$ and the discrete quantity computed using its detail coefficients d^j arising from the nonlinear and non-separable multiscale representation. We show that the existing results in one dimension naturally extend to our framework due to the caracterization of Besov spaces by mixed finite differences.

Lower estimates of the Besov norm are associated to a so-called *direct* theorem while upper estimates are associated to a so-called *inverse* theorem. Note that a similar technique was applied in [6] in a wavelet setting.

5.1. Direct Theorem

Let v be a function in some Besov space $B^s_{p,q}(\mathbb{R}^d)$ with $p, q \geq 1$ and $s > 0$, $(v^0, (d^j)_{j\geq 0})$ be its nonlinear multi-scale representation. We now show under what conditions we are able to get a lower estimate of $\|v\|_{B^s_{p,q}(\mathbb{R}^d)}$ using $(v^0, (d^j)_{j\geq 0})$. To prove such a result, we need to have first an estimate of the norm of the prediction error:

Lemma 5.1. *Assume that the quasi-linear prediction rule S(w) exactly reproduces polynomials of degree $N-1$ then the following estimation holds*

$$\|e^j\|_{\ell^p(\mathbb{Z}^d)} \lesssim \sum_{|n|=N} \|\nabla^n v^j\|_{\ell^p(\mathbb{Z}^d)}. \tag{17}$$

Proof: Let us compute

$$e^j_k(w) := v^j_k - \sum_{\|k-Ml\|_\infty \leq K} a_{k-Ml}(w) v^{j-1}_l.$$

Using (9), we can write it down as

$$\begin{aligned} e^j_k(w) &= v^j_k - m^{-1} \sum_{\substack{l\in\mathbb{Z}^d \\ \|k-Ml\|_\infty \leq K}} a_{k-Ml}(w) \sum_{\substack{n\in\mathbb{Z}^d \\ \|n-Ml\|_\infty \leq P}} \tilde{h}_{n-Ml} v^j_n \\ &= v^j_k - m^{-1} \sum_{\substack{n\in\mathbb{Z}^d \\ \|k-n\|_\infty \leq K+P}} v^j_n \sum_{\substack{l\in\mathbb{Z}^d \\ \|k-Ml\|_\infty \leq K}} a_{k-Ml}(w)\tilde{h}_{n-Ml} = \sum_{n\in F(k)} b_{k,n}(w) v^j_n, \end{aligned}$$

where $b_{k,n}(w) = \sum_{\substack{l\in\mathbb{Z}^d \\ \|k-Ml\|_\infty \leq K}} a_{k-Ml}(w)\tilde{h}_{n-Ml}$, and $F(k) = \{n \in \mathbb{Z}^d : \|n - k\|_\infty \leq P + K\}$ is a finite set for any given k. For any $k \in \mathbb{Z}^d$ let us define a vector $b_k(w) := (b_{k,n}(w))_{n\in F(k)}$. By hypothesis, $e^j(w) = 0$ if there exists $p \in \Pi_{N'}$, $0 \leq N' < N$ such that $v_k = p(k)$. Consequently, for any $q \in \mathbb{Z}^d, |q| < N$, $b_k(w)$ is orthogonal to any polynomial sequence associated to the polynomial $l^q = l_1^{q_1} \cdot \ldots \cdot l_d^{q_d}$, thus it can be written in terms of a basis of the space orthogonal to the space spanned by these vectors. According to [12], Theorem 4.3, we can take $\{\nabla^\mu \delta_{\cdot - l}, |\mu| = N, l \in \mathbb{Z}^d\}$ as a basis of this space. By denoting $c_l^\mu(w)$ the coordinates of $b_k(w)$ in this basis, we may write:

$$b_{k,n}(w) = \sum_{|\mu|=N} \sum_{l \in \mathbb{Z}^d} c_l^\mu(w) \nabla^\mu \delta_{n-l}$$

and taking $w = v^{j-1}$ we get

$$e_k^j := e_k^j(v^{j-1}) = \sum_{n \in F(k)} \sum_{|\mu|=N} \sum_{l \in \mathbb{Z}^d} c_l^\mu(v^{j-1}) \nabla^\mu \delta_{n-l} v_n^j = \sum_{n \in F(k)} \sum_{|\mu|=N} c_n^\mu(v^{j-1}) \nabla^\mu v_n^j. \tag{18}$$

Finally, we use (12) to conclude that the coefficients $b_{k,n}(v^{j-1})$ and $c_l^\mu(v^{j-1})$ are bounded independently of k, n and w, and (17) follows from (18). □

In what follows, we will use the definition of isotropic matrices:

Definition 5.1. *A matrix M is called isotropic if it is similar to the diagonal matrix $diag(\sigma_1, \ldots, \sigma_d)$, i.e. there exists an invertible matrix Λ such that*

$$M = \Lambda^{-1} diag(\sigma_1, \ldots, \sigma_d) \Lambda, \tag{19}$$

with $\sigma_1, \ldots, \sigma_d$ being the eigenvalues of matrix M, $\sigma := |\sigma_1| = \ldots = |\sigma_d| = m^{\frac{1}{d}}$.

Moreover, for any given norm in $\mathbb{R}^d$ there exist constants $C_1, C_2 > 0$ such that for any integer n and for any $v \in \mathbb{R}^d$

$$C_1 m^{\frac{n}{d}} \|v\| \leq \|M^n v\| \leq C_2 m^{\frac{n}{d}} \|v\|.$$

Lemma 17 enables us to compute the lower estimate:

Theorem 5.1. *If for $p, q \geq 1$ and some positive s, v belongs to $B^s_{p,q}(\mathbb{R}^d)$, if the quasi-linear prediction rule exactly reproduces polynomials of degree $N-1$ with $N > s$, if the matrix M is isotropic and if the equivalence* (8) *is satisfied, then*

$$\|v^0\|_{\ell^p(\mathbb{Z}^d)} + \|(m^{(s/d-1/p)j}\|(d^j_k)_{k\in\mathbb{Z}^d}\|_{\ell^p(\mathbb{Z}^d)})_{j\geq 0}\|_{\ell^q(\mathbb{Z}^d)} \lesssim \|v\|_{B^s_{p,q}(\mathbb{R}^d)}. \qquad (20)$$

Proof: Using the Hölder inequality and the fact that $\tilde{\varphi}$ is compactly supported, we first obtain

$$\|v^0\|_{\ell^p(\mathbb{Z}^d)} = \|(\langle v, \tilde{\varphi}(\cdot - k)\rangle)_{k\in\mathbb{Z}^d}\|_{\ell^p(\mathbb{Z}^d)} \lesssim \|(\|v\|_{L^p(\mathrm{Supp}(\tilde{\varphi}(\cdot-k)))})_{k\in\mathbb{Z}^d}\|_{\ell^p(\mathbb{Z}^d)} \lesssim \|v\|_{L^p(\mathbb{R}^d)}.$$

Let us then consider a *quasi-linear* prediction rule which exactly reproduces polynomials of degree $N-1$. Since $\|e^j_\cdot\|_{\ell^p(\mathbb{Z}^d)} \sim \|d^{j-1}_\cdot\|_{\ell^p(\mathbb{Z}^d)}$, by Lemma 5.1 we get

$$\|(m^{(s/d-1/p)j}\|(d^j_\cdot)\|_{\ell^p(\mathbb{Z}^d)})_{j\geq 0}\|_{\ell^q(\mathbb{Z}^d)} \lesssim \|(m^{(s/d-1/p)j}\sum_{|n|=N}\|(\nabla^n v^j_\cdot)\|_{\ell^p(\mathbb{Z}^d)})_{j\geq 0}\|_{\ell^q(\mathbb{Z}^d)}.$$

We have successively

$$\begin{aligned}
\sum_{|n|=N}\|\nabla^n v^j\|_{\ell^p(\mathbb{Z}^d)} &= \sum_{|n|=N}\|\nabla^n(\langle v, m^j\tilde{\varphi}(M^j\cdot -k)\rangle)_{k\in\mathbb{Z}^d}\|_{\ell^p(\mathbb{Z}^d)} \\
&= \sum_{|n|=N}\|((\langle \nabla^n_{M^{-j}} v, m^j\tilde{\varphi}(M^j\cdot -k)\rangle)_{k\in\mathbb{Z}^d}\|_{\ell^p(\mathbb{Z}^d)} \\
&\lesssim m^{j/p}\|\sum_{|n|=N}(\|\nabla^n_{M^{-j}} v\|_{L^p(\mathrm{Supp}(\tilde{\varphi}(M^j\cdot-k)))})_{k\in\mathbb{Z}^d}\|_{\ell^p(\mathbb{Z}^d)} \\
&\lesssim m^{j/p}\sum_{|n|=N}\|\nabla^n_{M^{-j}} v\|_{L^p(\mathbb{R}^d)} \\
&\lesssim m^{j/p}\tilde{\omega}_N(v, C_2 m^{-j/d})_{L^p},
\end{aligned}$$

since M is isotropic. Furthermore, for any integer $C > 0$ and any $t > 0$, $\tilde{\omega}_N(v, Ct)_{L^p} \leq C\tilde{\omega}_N(v, t)_{L^p}$. Thus,

$$\sum_{|n|=N} \|\nabla^n v^j\|_{\ell^p(\mathbb{Z}^d)} \lesssim m^{j/p} \tilde{\omega}_N(v, m^{-j/d})_{L^p},$$

which implies (20). □

5.2. Inverse Theorems

We consider the sequence $(v^0, (d^j)_{j\geq 0})$ and we study the convergence of the reconstruction process:

$$v^j = S(v^{j-1})v^{j-1} + Ed^{j-1},$$

throughout the study of the limit of the sequence of functions

$$v_j(x) = \sum_{k\in\mathbb{Z}^d} v^j_k \varphi(M^j x - k), \tag{21}$$

where φ is defined in (1). More precisely, we show that under certain conditions on the sequence $(v^0, (d^j)_{j\geq 0})$ and on φ, v_j converges to some function v belonging to a Besov space.

For that purpose, we establish that if the *quasi-linear* prediction rule $S(w)$ reproduces polynomials of degree $N-1$ then all the mixed finite differences of order lesser than N can be defined using difference operators:

Proposition 5.1. *Let $S(w)$ be a quasi-linear prediction rule reproducing polynomials of degree $N-1$. Then for any l, $0 < l \leq N$ there exists a difference operator S_l such that:*

$$\Delta^l S(w)u := S_l(w)\Delta^l u,$$

for all $u, w \in \ell^\infty(\mathbb{Z}^d)$.

The proof is available in [18], Proposition 1. In contrast to the tensor product case studied in [17], the operator $S_l(w)$ is multi-dimensional and is defined from $(\ell^\infty(\mathbb{Z}^d))^{q_l}$ onto $(\ell^\infty(\mathbb{Z}^d))^{q_l}$, $q_l = \#\{\mu, |\mu| = l\}$, and cannot be reduced to a set of difference operators in some given directions.

The inverse theorem proved in this section is based on the contractivity of the difference operators. This is done by studying the joint spectral radius, which we now define:

Definition 5.2. *Let us consider a set of difference operators $(S_l)_{l\geq 0}$, defined in Proposition 5.1 with $S_0 := S$. The joint spectral radius in $(\ell^p(\mathbb{Z}^d))^{q_l}$ of S_l is given by*

$$\begin{aligned}\rho_p(S_l) &:= \inf_{j>0} \sup_{(w^0,\cdots,w^{j-1})\in(\ell^p(\mathbb{Z}^d))^j} \|S_l(w^{j-1})\cdot\ldots\cdot S_l(w^0)\|^{1/j}_{(\ell^p(\mathbb{Z}^d))^{q_l}\to(\ell^p(\mathbb{Z}^d))^{q_l}} \\ &= \inf_{j>0}\{\rho, \|S_l(w^{j-1})\cdots S_l(w^0)\Delta^l v\|_{(\ell^p(\mathbb{Z}^d))^{q_l}} \lesssim \rho^j\|\Delta^l v\|_{(\ell^p(\mathbb{Z}^d))^{q_l}}, \forall v\in\ell^p(\mathbb{Z}^d)\}.\end{aligned}$$

Remark 5.1. *When $v^j = S(v^{j-1})v^{j-1}$, for all $j > 0$ we may write:*

$$\Delta^l S(v^j)v^j = S_l(S(v^{j-1})v^{j-1})\Delta^l v^j = S_l(S(v^{j-1})v^{j-1})S_l(v^{j-1})\Delta^l v^{j-1} = \cdots := (S_l)^j v^0.$$

This naturally leads to another definition of the joint spectral radius by putting $w^j = S^j v^0$ in the above definition. In [19], the following definition was introduced to study the convergence and stability of one-dimensional power-P scheme. In that context, the joint spectral radius in $(\ell^p(\mathbb{Z}^d))^{q_l}$ of S_l is changed into

$$\begin{aligned}\tilde\rho_p(S_l) &:= \inf_{j>0} \|(S_l)^j\|^{1/j}_{(\ell^p(\mathbb{Z}^d))^{q_l}\to(\ell^p(\mathbb{Z}^d))^{q_l}} \\ &= \inf_{j>0}\{\rho, \|\Delta^l S^j v\|_{(\ell^p(\mathbb{Z}^d))^{q_l}} \lesssim \rho^j\|\Delta^l v\|_{(\ell^p(\mathbb{Z}^d))^{q_l}}, \forall v\in\ell^p(\mathbb{Z}^d)\}.\end{aligned}$$

Since our prediction operator is quasi-linear, the definition (5.2) *is more appropriate.*

Before we prove the *inverse* theorem, we need to establish some extensions to the non-separable case of results obtained in [17]:

Lemma 5.2. *Let $S(w)$ be a quasi-linear prediction rule that exactly reproduces polynomials of degree $N-1$. Then,*

$$\|v_{j+1} - v_j\|_{L^p(\mathbb{R}^d)} \lesssim m^{-j/p}\left(\|\Delta^N v^j\|_{(\ell^p(\mathbb{Z}^d))^{q_N}} + \|d^j\|_{\ell^p(\mathbb{Z}^d)}\right). \tag{22}$$

Moreover, for any $\rho_p(S_N) < \rho < m^{1/p}$ there exists an n such that,

$$m^{-j/p}\|\Delta^N v^j\|_{(\ell^p(\mathbb{Z}^d))^{q_N}} \lesssim \delta^j\|v^0\|_{\ell^p(\mathbb{Z}^d)} + \sum_{r=0}^{t-1}\delta^{nr}\sum_{l=j-(r+1)n+1}^{l=j-rn} m^{-l/p}\|d^{l-1}\|_{\ell^p(\mathbb{Z}^d)} \tag{23}$$

where $\delta = \rho m^{-1/p}$ and $t = \lfloor j/n \rfloor$.

Proof: Using the definition of functions $v_j(x)$ and the scaling equation (1), we get

$$\begin{aligned} v_{j+1}(x) - v_j(x) &= \sum_k v_k^{j+1}\varphi(M^{j+1}x - k) - \sum_k v_k^j\varphi(M^j x - k) \\ &= \sum_k ((S(v^j)v^j)_k + d_k^j)\varphi(M^{j+1}x - k) - \sum_k v_k^j \sum_l g_{l-Mk}\varphi(M^{j+1}x - l) \\ &= \sum_k ((S(v^j)v^j)_k - \sum_l g_{k-Ml}v_l^j)\varphi(M^{j+1}x - k) + \sum_k d_k^j\varphi(M^{j+1}x - k). \end{aligned}$$

If $S(w)$ exactly reproduces polynomials of degree $N-1$ for all w, then having in mind that

$$Sv^j := \sum_l g_{k-Ml}v_l^j, \tag{24}$$

is the linear prediction which exactly reproduces polynomials of degree $N-1$ and using the same arguments as in Lemma 5.1, we get

$$\|\sum_k ((S(v^j)v^j)_k - Sv_k^j)\varphi(M^{j+1}x-k)\|_{L^p(\mathbb{R}^d)} \lesssim m^{-j/p}\|\Delta^N v^j\|_{(\ell^p(\mathbb{Z}^d))^{q_N}}. \tag{25}$$

The proof of (22) is thus complete. To prove (23), we note that for any $\rho_p(S_N) < \rho < m^{1/p}$, there exists an n such that for all v:

$$\|(S_N)^n v\|_{(\ell^p(\mathbb{Z}^d))^{q_N}} \leq \rho^n \|v\|_{(\ell^p(\mathbb{Z}^d))^{q_N}}. \tag{26}$$

Using the boundedness of the operator S_N, we obtain:

$$\begin{aligned}
\|\Delta^N v^n\|_{(\ell^p(\mathbb{Z}^d))^{q_N}} &\leq \|S_N \Delta^N v^{n-1}\|_{(\ell^p(\mathbb{Z}^d))^{q_N}} + \|\Delta^N d^{n-1}\|_{(\ell^p(\mathbb{Z}^d))^{q_N}} \\
&\leq \|(S_N)^n \Delta^N v^0\|_{(\ell^p(\mathbb{Z}^d))^{q_N}} + D \sum_{l=0}^{n-1} \|d^l\|_{\ell^p(\mathbb{Z}^d)} \\
&\leq \rho^n \|\Delta^N v^0\|_{(\ell^p(\mathbb{Z}^d))^{q_N}} + D \sum_{l=0}^{n-1} \|d^l\|_{\ell^p(\mathbb{Z}^d)}
\end{aligned}$$

Then for any j, define $t := \lfloor j/n \rfloor$, after t iterations of the above inequality, we get:

$$\|\Delta^N v^j\|_{(\ell^p(\mathbb{Z}^d))^{q_N}} \leq \rho^{nt} \|\Delta^N v^{j-nt}\|_{(\ell^p(\mathbb{Z}^d))^{q_N}} + D \sum_{r=0}^{t-1} \rho^{nr} \sum_{l=rn}^{(r+1)n-1} \|d^{j-1-l}\|_{\ell^p(\mathbb{Z}^d)}$$

Then putting as in [10], $\delta = \rho m^{-1/p}$, and $A_j = m^{-j/p} \|\Delta^N v^j\|_{(\ell^p(\mathbb{Z}^d))^{r_N^d}}$, we get:

$$A_j \leq \delta^{nt} A_{j-nt} + D \sum_{r=0}^{t-1} \delta^{nr} \sum_{l=nr}^{l=(r+1)n-1} m^{-(j-l)/p} \|d^{j-1-l}\|_{\ell^p(\mathbb{Z}^d)}$$

Then, we may write, due the boundedness of $S^{(k)}$, for $j' < n$:

$$A_{j'} \lesssim \|v_0\|_{\ell^p(\mathbb{Z}^d)} + \sum_{l=1}^{j'} m^{-l/p} \|d^{l-1}\|_{\ell^p(\mathbb{Z}^d)}$$

which finally leads to:

$$m^{-j/p} \|\Delta^N v^j\|_{(\ell^p(\mathbb{Z}^d))^{r_N^d}} \lesssim \delta^j \|v^0\|_{\ell^p(\mathbb{Z}^d)} + \sum_{r=0}^{t-1} \delta^{nr} \sum_{l=j-n(r+1)+1}^{l=j-nr} m^{-l/p} \|d^{l-1}\|_{\ell^p(\mathbb{Z}^d)}$$

$\square$

By abusing a little bit terminology, we say that φ exactly reproduces polynomials if the underlying subdivision scheme does. With this in mind, we are ready to state the *inverse* theorems: the first one deals with L^p convergence under the main hypothesis $\rho_p(S_1) < m^{1/p}$, while the second deals with the convergence in $B^s_{p,q}(\mathbb{R}^d)$ under the main hypothesis $\rho_p(S_N) < m^{1/p-s/d}$ for some $N > 1$ and $N-1 \leq s < N$.

Theorem 5.2. *Let $S(w)$ be a quasi-linear prediction rule reproducing the constants. Assume that $\rho_p(S_1) < m^{1/p}$, if*

$$\|v^0\|_{\ell^p(\mathbb{Z}^d)} + \sum_{j\geq 0} m^{-j/p}\|d^j\|_{\ell^p(\mathbb{Z}^d)} < \infty,$$

then the limit function v belongs to $L^p(\mathbb{R}^d)$ and

$$\|v\|_{L^p(\mathbb{R}^d)} \leq \|v^0\|_{\ell^p(\mathbb{Z}^d)} + \sum_{j\geq 0} m^{-j/p}\|d^j\|_{\ell^p(\mathbb{Z}^d)}. \tag{27}$$

Proof: From estimates (22) and (23), for any $\rho_p(S_1) < \rho < m^{1/p}$ there exists an n such that:

$$\|v_{j+1} - v_j\|_{L^p(\mathbb{R}^d)} \lesssim \delta^j\|v^0\|_{\ell^p(\mathbb{Z}^d)} + \sum_{r=0}^{s} \delta^{nr} \sum_{l=j-n(r+1)+1}^{l=j-nr} m^{-l/p}\|d^{l-1}\|_{\ell^p(\mathbb{Z}^d)} + m^{-j/p}\|d^j\|_{\ell^p(\mathbb{Z}^d)},$$

from which we deduce that:

$$\|v\|_{L^p(\mathbb{R}^d)} \le \|v_0\|_{L^p(\mathbb{R}^d)} + \sum_{j\ge 0} \|v_{j+1} - v_j\|_{L^p(\mathbb{R}^d)} \tag{28}$$

$$\lesssim \|v^0\|_{\ell^p(\mathbb{Z}^d)} + \sum_{j\ge 0} \delta^j \|v^0\|_{\ell^p(\mathbb{Z}^d)} + \sum_{r=0}^{t-1} \delta^{nr} \sum_{l=j-n(r+1)+1}^{l=j-nr} m^{-l/p} \|d^{l-1}\|_{\ell^p(\mathbb{Z}^d)} + m^{-j/p} \|d^j\|_{\ell^p(\mathbb{Z}^d)}$$

$$\lesssim \|v^0\|_{\ell^p(\mathbb{Z}^d)} + \sum_{t=0}^{\infty} \sum_{q=0}^{n-1} \sum_{r'=1}^{t} \delta^{n(t-r')} \sum_{l=r'n-n+q+1}^{l=r'n+q} m^{-l/p} \|d^{l-1}\|_{\ell^p(\mathbb{Z}^d)} + \sum_{j>0} m^{-j/p} \|d^j\|_{\ell^p(\mathbb{Z}^d)}$$

$$\lesssim \|v^0\|_{\ell^p(\mathbb{Z}^d)} + \sum_{r'=1}^{\infty} \sum_{t>r'} \delta^{n(t-r')} \sum_{q=0}^{n-1} \sum_{l=r'n-n+q+1}^{l=r'n+q} m^{-l/p} \|d^{l-1}\|_{\ell^p(\mathbb{Z}^d)} + \sum_{j>0} m^{-j/p} \|d^j\|_{\ell^p(\mathbb{Z}^d)}$$

$$\lesssim \|v^0\|_{\ell^p(\mathbb{Z}^d)} + \sum_{j>0}^{\infty} m^{-j/p} \|d^{j-1}\|_{\ell^p(\mathbb{Z}^d)} \tag{29}$$

The last equality being obtained remarking that $\sum\limits_{t>r'} \delta^{n(t-r')} = \frac{1}{1-\delta^n}$. This proves (27). □

Now, we extend this result to the case of Besov spaces:

Theorem 5.3. *Let $S(w)$ be a quasi-linear prediction rule exactly reproducing polynomials of degree $N-1$ and let φ exactly reproduce polynomials of degree $N-1$. Assume that $\rho_p(S_N) < m^{1/p-s/d}$ for some $N > s \ge N-1$. If $(v^0, d^0, d^1, \ldots)$ are such that*

$$\|v^0\|_{\ell^p(\mathbb{Z}^d)} + \|(m^{(s/d-1/p)j} \|(d_k^j)_{k\in\mathbb{Z}^d}\|_{\ell^p(\mathbb{Z}^d)})_{j\ge 0}\|_{\ell^q(\mathbb{Z}^d)} < \infty,$$

the limit function v belongs to $B^s_{p,q}(\mathbb{R}^d)$ and

$$\|v\|_{B^s_{p,q}(\mathbb{R}^d)} \lesssim \|v^0\|_{\ell^p(\mathbb{Z}^d)} + \|(m^{(s/d-1/p)j} \|(d_k^j)_{k\in\mathbb{Z}^d}\|_{\ell^p(\mathbb{Z}^d)})_{j\ge 0}\|_{\ell^q(\mathbb{Z}^d)}. \tag{30}$$

Proof: First, by Hölder inequality for any $q, q' > 0, \frac{1}{q} + \frac{1}{q'} = 1$, it holds that

$$\sum_{l\geq 0}\|d^l\|_{\ell^p(\mathbb{Z}^d)}m^{-l/p} \leq \|(m^{(s/d-1/p)j}\|(d_k^j)_{k\in\mathbb{Z}^d}\|_{\ell^p(\mathbb{Z}^d)})_{j\geq 0}\|_{\ell^q(\mathbb{Z}^d)}\|(m^{-js/d})_{j\geq 0}\|_{\ell^{q'}(\mathbb{R}^d)}$$

$$\lesssim \|(m^{(s/d-1/p)j}\|(d_k^j)_{k\in\mathbb{Z}^d}\|_{\ell^p(\mathbb{Z}^d)})_{j\geq 0}\|_{\ell^q(\mathbb{Z}^d)},$$

and finally,

$$\|v\|_{L^p(\mathbb{R}^d)} \lesssim \|v^0\|_{\ell^p(\mathbb{Z}^d)} + \|(m^{(s/d-1/p)j}\|(d_k^j)_{k\in\mathbb{Z}^d}\|_{\ell^p(\mathbb{Z}^d)})_{j\geq 0}\|_{\ell^q(\mathbb{Z}^d)}.$$

It remains to evaluate the semi-norm $|v|_{B^s_{p,q}(\mathbb{R}^d)} := \|(m^{js/d}\tilde{\omega}_N(v,m^{-j/d})_{L^p})_{j\geq 0}\|_{\ell^q(\mathbb{Z}^d)}$. For each $j\geq 0$, we have

$$\tilde{\omega}_N(v,m^{-j/d})_{L^p} \leq \tilde{\omega}_N(v-v_j,m^{-j/d})_{L^p} + \tilde{\omega}_N(v_j,m^{-j/d})_{L^p}. \tag{31}$$

Note that the property (23) can be extended to the case where $\rho_p(S_N) < m^{1/p-s/d}$. Making the same kind of computation as in the proof of (23), one can prove that for any $\rho_p(S_N) < \rho < m^{1/p-s/d}$ there exists an n such that:

$$m^{-j(1/p-s/d)}\|\Delta^N v^j\|_{(\ell^p(\mathbb{Z}^d))^{q_N}} \lesssim \delta^j\|v^0\|_{\ell^p(\mathbb{Z}^d)} + \sum_{r=0}^{t-1}\delta^{nr}\sum_{l=j-(r+1)n+1}^{l=j-rn} m^{-l(1/p-s/d)}\|d^{l-1}\|_{\ell^p(\mathbb{Z}^d)}$$

where $\delta := \rho m^{-1/p+s/d}$ and $t = \lfloor j/n \rfloor$. Then, we can deduce that:

$$\begin{aligned}
\|\Delta^N v^j\|_{(\ell^p(\mathbb{Z}^d))^{q_N}} &\lesssim \rho^j(\|v^0\|_{\ell^p(\mathbb{Z}^d)} + \delta^{-j}\sum_{r=0}^{s-1}\delta^{nr}\sum_{l=j-(r+1)n+1}^{l=j-rn} m^{-l(1/p-s/d)}\|d^{l-1}\|_{\ell^p(\mathbb{Z}^d)}) \\
&\lesssim \rho^j(\|v^0\|_{\ell^p(\mathbb{Z}^d)} + \sum_{r=0}^{s-1}\delta^{-n+1}\sum_{l=j-(r+1)n+1}^{l=j-rn} \rho^{-l}\|d^{l-1}\|_{\ell^p(\mathbb{Z}^d)}) \\
&\lesssim \rho^j(\|v^0\|_{\ell^p(\mathbb{Z}^d)} + \sum_{l=0}^{j}\rho^{-l}\|d^{l-1}\|_{\ell^p(\mathbb{Z}^d)}) \qquad (32)
\end{aligned}$$

For the first term on the right hand side of (31), one has using (32):

$$\begin{aligned}
\tilde{\omega}_N(v - v_j, m^{-j/d})_{L^p} &\lesssim \sum_{l\geq j} \|v_{l+1} - v_l\|_{L^p(\mathbb{R}^d)} \\
&\lesssim \sum_{l\geq j} m^{-l/p}(\|\Delta^N v^l\|_{(\ell^p(\mathbb{Z}^d))^{q_N}} + \|d^l\|_{\ell^p(\mathbb{Z}^d)}) \\
&\lesssim \sum_{l\geq j} m^{-l/p}\left(\rho^l\|v^0\|_{\ell^p(\mathbb{Z}^d)} + \sum_{k=0}^{l} \rho^{l-k}\|d^k\|_{\ell^p(\mathbb{Z}^d)}\right).
\end{aligned}$$

For the first term, choosing since $\rho < m^{1/p}$, we have

$$\sum_{l\geq j} m^{-l/p}\rho^l\|v^0\|_{\ell^p(\mathbb{Z}^d)} \sim m^{-j/p}\rho^j\|v^0\|_{\ell^p(\mathbb{Z}^d)}.$$

For the second term, we get

$$\begin{aligned}
&\sum_{l\geq j} m^{-l/p}\sum_{k=0}^{l}\rho^{l-k}\|d^k\|_{\ell^p(\mathbb{Z}^d)} \\
&\quad = m^{-j/p}\sum_{k=0}^{j}\rho^{j-k}\|d^k\|_{\ell^p(\mathbb{Z}^d)} + \sum_{l>j} m^{-l/p}\sum_{k=0}^{l}\rho^{l-k}\|d^k\|_{\ell^p(\mathbb{Z}^d)} \\
&\quad \lesssim m^{-j/p}\sum_{k=0}^{j}\rho^{j-k}\|d^k\|_{\ell^p(\mathbb{Z}^d)} + \sum_{k\geq 0}\sum_{l>\max(k,j)} m^{-l/p}\rho^{l-k}\|d^k\|_{\ell^p(\mathbb{Z}^d)} \\
&\quad \lesssim m^{-j/p}\sum_{k=0}^{j}\rho^{j-k}\|d^k\|_{\ell^p(\mathbb{Z}^d)} + \sum_{k\geq 0}\|d^k\|_{\ell^p(\mathbb{Z}^d)}\rho^{-k}\sum_{l>\max(k,j)}\rho^l m^{-l/p} \\
&\quad \lesssim m^{-j/p}\sum_{k=0}^{j}\rho^{j-k}\|d^k\|_{\ell^p(\mathbb{Z}^d)} + \sum_{k>j} m^{-k/p}\|d^k\|_{\ell^p(\mathbb{Z}^d)}.
\end{aligned}$$

Similarly, for the second term on the right hand side of (31), one has

$$\tilde{\omega}_N(v_j, m^{-j/d})_{L^p} \lesssim \|v_j\|_{L^p(\mathbb{R}^d)} \lesssim \|v\|_{L^p(\mathbb{R}^d)}$$

The estimate of the semi-norm $|v|_{B^s_{p,q}}$ is then reduced to the estimates of $\|(m^{js/d}a_j)_{j\geq 0}\|_{\ell^q(\mathbb{Z}^d)}$, $\|(m^{js/d}b_j)_{j\geq 0}\|_{\ell^q(\mathbb{Z}^d)}$ and $\|(m^{js/d}c_j)_{j\geq 0}\|_{\ell^q(\mathbb{Z}^d)}$, with

$$\begin{aligned}
a_j &:= m^{-j/p}\rho^j \|v^0\|_{\ell^p(\mathbb{Z}^d)},\\
b_j &:= m^{-j/p}\rho^j \sum_{l=0}^{j} \rho^{-l}\|d^l\|_{\ell^p(\mathbb{Z}^d)},\\
c_j &:= \sum_{l>j} m^{-l/p}\|d^l\|_{\ell^p(\mathbb{Z}^d)}.
\end{aligned}$$

Recalling $\delta = m^{s/d-1/p}\rho < 1$, we write:

$$\|(\sigma^{js}a_j)_{j\geq 0}\|_{\ell^q(\mathbb{Z}^d)} = \|v^0\|_{\ell^p(\mathbb{Z}^d)}\|(\delta^j)_{j\geq 0}\|_{\ell^q(\mathbb{Z}^d)} \lesssim \|v^0\|_{\ell^p(\mathbb{Z}^d)}. \tag{33}$$

In order to estimate $\|(m^{js/d}b_j)_{j\geq 0}\|_{\ell^q(\mathbb{Z}^d)}$, we rewrite it in the following form:

$$\begin{aligned}
\|(m^{js/d}b_j)_{j\geq 0}\|_{\ell^q(\mathbb{Z}^d)} &= \|(m^{j(s/d-1/p)}\rho^j \sum_{l=0}^{j} \rho^{-l}\|d^l\|_{\ell^p})_{j\geq 0}\|_{\ell^q(\mathbb{Z}^d)}\\
&= \|(\delta^j \sum_{l=0}^{j} \delta^{-l} m^{(s/d-1/p)l}\|d^l\|_{\ell^p(\mathbb{Z}^d)})_{j\geq 0}\|_{\ell^q(\mathbb{Z}^d)}.
\end{aligned}$$

We, now, make use of the following discrete Hardy inequality: if $0<\delta<1$, then

$$\|(\delta^j \sum_{l=0}^{j} \delta^{-l}x_l)_{j\geq 0}\|_{\ell^q(\mathbb{Z}^d)} \lesssim \|(x_j)_{j\geq 0}\|_{\ell^q(\mathbb{Z}^d)}.$$

Applying it to $x_l = m^{(s/d-1/p)l}\|d^l\|_{\ell^p(\mathbb{Z}^d)}$ yields

$$\|(m^{js/d}b_j)_{j\geq 0}\|_{\ell^q(\mathbb{Z}^d)} \lesssim \|(m^{(s/d-1/p)j}\|(d^j_k)_{k\in\mathbb{Z}^d}\|_{\ell^p(\mathbb{Z}^d)})_{j\geq 0}\|_{\ell^q(\mathbb{Z}^d)}. \tag{34}$$

To estimate $\|(m^{js/d}c_j)_{j\geq 0}\|_{\ell^q(\mathbb{Z}^d)}$, we rewrite it as follows

$$\|(m^{js/d}c_j)_{j\geq 0}\|_{\ell^q(\mathbb{Z}^d)} \quad = \quad \|m^{js/d}\sum_{l>j} m^{-ls/d}\left(m^{l(s/d-1/p)}\|(d^l_k)\|_{\ell^p(\mathbb{Z}^d)}\right)\|_{\ell^q(\mathbb{Z}^d)}$$

and make use of another discrete Hardy inequality: if $\beta > 1$, then

$$\|(\beta^j \sum_{l>j} \beta^{-l} y_l)_{j\geq 0}\|_{\ell^q(\mathbb{Z}^d)} \lesssim \|(y_j)_{j\geq 0}\|_{\ell^q(\mathbb{Z}^d)}.$$

Taking $y_l = m^{l(s/d-1/p)}\|d^l\|_{\ell^p(\mathbb{Z}^d)}$, we obtain, since $s > N-1$,

$$\|(m^{js/d}c_j)_{j\geq 0}\|_{\ell^q(\mathbb{Z}^d)} \lesssim \|(m^{j(s/d-1/p)}\|(d^j_k)_{k\in\mathbb{Z}^d}\|_{\ell^p(\mathbb{Z}^d)})_{j\geq 0}\|_{\ell^q(\mathbb{Z}^d)}. \quad (35)$$

Then (30) follows by combining (33), (34) and (35). □

6. Stability of the Multi-Scale Representation

Here, we consider two data sets $(v^0, d^0, d^1, \ldots)$ and $(\tilde{v}^0, \tilde{d}^0, \tilde{d}^1, \ldots)$ corresponding to two reconstruction processes

$$v^j = S(v^{j-1})v^{j-1} + e^j = S(v^{j-1})v^{j-1} + Ed^{j-1} \quad (36)$$

and

$$\tilde{v}^j = S(\tilde{v}^{j-1})\tilde{v}^{j-1} + \tilde{e}^j = S(\tilde{v}^{j-1})\tilde{v}^{j-1} + E\tilde{d}^{j-1}. \quad (37)$$

where E is the matrix corresponding to the basis of the kernel of the projection operator. In that context, v is the limit of $v_j(x) = \sum_{k\in\mathbb{Z}^d} v^j_k \varphi(M^j x - k)$ (and similarly for $\tilde{v}$).

To study the stability of the multi-scale representation in $L^p(\mathbb{R}^d)$, we need the following Lemma:

Lemma 6.1. *Let $S(w)$ be a quasi-linear prediction rule that exactly reproduces polynomials of degree $N-1$. Then, putting $u_j = v_j - v_{j-1}$ and $\tilde{u}_j = \tilde{v}_j - \tilde{v}_{j-1}$ we get*

$$\|u_j - \tilde{u}_j\|_{L^p(\mathbb{R}^d)} \lesssim m^{-j/p}\left(\|\Delta^N(v^{j-1} - \tilde{v}^{j-1})\|_{(\ell^p(\mathbb{Z}^d))^{q_N}} + \|d^{j-1} - \tilde{d}^{j-1}\|_{\ell^p(\mathbb{Z}^d)}\right). \quad (38)$$

Proof: By definition of the linear prediction operator S, see (24), we can write

$$\begin{aligned}
\|u_j - \tilde{u}_j\|_{L^p(\mathbb{R}^d)} &\leq m^{-j/p}\|(S(v_{j-1}) - S)v^{j-1} + d^{j-1} - ((S(\tilde{v}_{j-1}) - S)\tilde{v}^{j-1} + \tilde{d}^{j-1})\|_{\ell^p(\mathbb{Z}^d)} \\
&\leq m^{-j/p}\left(\|\Delta^N(v^{j-1} - \tilde{v}^{j-1})\|_{(\ell^p(\mathbb{Z}^d))^{q_N}} + \|d^{j-1} - \tilde{d}^{j-1}\|_{\ell^p(\mathbb{Z}^d)}\right).
\end{aligned}$$

□

Now, we study the stability of the multi-scale representation in $L^p(\mathbb{R}^d)$, which is stated in the following result:

Theorem 6.1. *Let $S(w)$ be a quasi-linear prediction rule that reproduces the constants and suppose that there exist a $\rho < m^{1/p}$ and and an n such that*

$$\|(S_1)^n w - (S_1)^n v\|_{\ell^p(\mathbb{Z}^d)^d} \leq \rho^n \|v - w\|_{\ell^p(\mathbb{Z}^d)^d}. \quad \forall v, w \in \ell^p(\mathbb{Z}^d)^d \tag{39}$$

Assume that v_j and $\tilde{v}_j$ converge to v and $\tilde{v}$ in $L^p(\mathbb{R}^d)$ then we have:

$$\|v - \tilde{v}\|_{L^p(\mathbb{R}^d)} \lesssim \|v^0 - \tilde{v}^0\|_{\ell^p(\mathbb{Z}^d)} + \sum_{l \geq 0} m^{-l/p}\|d^l - \tilde{d}^l\|_{\ell^p(\mathbb{Z}^d)}. \tag{40}$$

Proof: First remark that due to the hypothesis (39) and the fact that the fact that S_1 is bounded, we may write that

$$\begin{aligned}
\|\Delta^1(v^n - \tilde{v}^n)\|_{(\ell^p(\mathbb{Z}^d))^d} &\leq \|S_1\Delta^1 v^{n-1} - S_1\Delta^1\tilde{v}^{n-1}\|_{(\ell^p(\mathbb{Z}^d))^d} + \|\Delta^1(d^{n-1} - \tilde{d}^{n-1})\|_{(\ell^p(\mathbb{Z}^d))^d} \\
&\leq \|(S_1)^n\Delta^1 v^0 - (S_1)^n\Delta^1\tilde{v}^0\|_{(\ell^p(\mathbb{Z}^d))^d} + D\sum_{l=0}^{n-1}\|d^l - \tilde{d}^l\|_{\ell^p(\mathbb{Z}^d)} \\
&\leq \rho^n\|\Delta^1 v^0 - \Delta^1\tilde{v}^0\|_{(\ell^p(\mathbb{Z}^d))^d} + D\sum_{l=0}^{n-1}\|d^l - \tilde{d}^l\|_{\ell^p(\mathbb{Z}^d)}
\end{aligned}$$

Then using the same kind of arguments as in the proof of (23), we can write:

$$m^{-j/p}\|\Delta^1(v^j - \tilde{v}^j)\|_{(\ell^p(\mathbb{Z}^d))^d} \lesssim \delta^j\|v^0 - \tilde{v}^0\|_{\ell^p(\mathbb{Z}^d)} + \sum_{r=0}^{t-1}\delta^{nr}\sum_{l=j-n(r+1)+1}^{l=j-nr} m^{-l/p}\|d^{l-1} - \tilde{d}^{l-1}\|_{\ell^p(\mathbb{Z}^d)}. \tag{4}$$

Now by Lemma 6.1, relation (41) and by then using the same kind of argument as in (28), one has:

$$\begin{aligned}\|v-\tilde{v}\|_{L^p(\mathbb{R}^d)} &\leq \sum_{j>0}\|u_j-\tilde{u}_j\|_{L^p(\mathbb{R}^d)}+\|v_0-\tilde{v}_0\|_{L^p(\mathbb{R}^d)}\\ &\leq \sum_{j>0} m^{-j/p}\left(\|\Delta^1(v^{j-1}-\tilde{v}^{j-1})\|_{(\ell^p(\mathbb{Z}^d))^d}+\|d^{j-1}-\tilde{d}^{j-1}\|_{\ell^p(\mathbb{Z}^d)}\right)+\|v_0-\tilde{v}_0\|_{\ell^p(\mathbb{Z}^d)}\\ &\leq \|v^0-\tilde{v}^0\|_{\ell^p(\mathbb{Z}^d)}+\sum_{l=0}^{j} m^{-l/p}\|d^{l-1}-\tilde{d}^{l-1}\|_{\ell^p(\mathbb{Z}^d)}.\end{aligned}$$

□

In view of the inverse inequality (30), it seems natural to seek an inequality of type

$$\|v-\tilde{v}\|_{B^s_{p,q}(\mathbb{R}^d)} \lesssim \|v^0-\tilde{v}^0\|_{\ell^p(\mathbb{Z}^d)}+\|(m^{(s/d-1/p)j}\|d^j_{.}-\tilde{d}^j_{.}\|_{\ell^p(\mathbb{Z}^d)})_{j\geq 0}\|_{\ell^q(\mathbb{Z}^d)}. \quad (42)$$

We now prove a stability theorem in Besov space $B^s_{p,q}(\mathbb{R}^d)$ in the following theorem:

Theorem 6.2. *Let $S(w)$ be a quasi-linear prediction rule which exactly reproduces polynomials of degree $N-1$. Assume that v_j and $\tilde{v}_j$ converge to v and $\tilde{v}$ in $B^s_{p,q}(\mathbb{R}^d)$ respectively and that there exist $\rho < m^{1/p-s/d}$ and an n such that:*

$$\|(S_N)^n w-(S_N)^n v\|_{(\ell^p(\mathbb{Z}^d))^{q_N}} \leq \rho^n\|w-v\|_{(\ell^p(\mathbb{Z}^d))^{q_N}} \quad \forall v,w\in(\ell^p(\mathbb{Z}^d))^{q_N}. \quad (43)$$

Then, we get that:

$$\|v-\tilde{v}\|_{B^s_{p,q}(\mathbb{R}^d)} \lesssim \|v^0-\tilde{v}^0\|_{\ell^p(\mathbb{Z}^d)}+\|(m^{-j(1/p-s/d)}\|(d^j_k-\tilde{d}^j_k)_{k\in\mathbb{Z}^d}\|_{\ell^p(\mathbb{Z}^d)})_{j\geq 0}\|_{\ell^q(\mathbb{Z}^d)}. \quad (44)$$

Remark 6.1. *We shall note that a similar study was proposed in the one-dimensional case to study the stability of the multiscale representation based on so-called r-shift invariant subdivision operators [19]. In that paper, the stability is obtained when $\rho_p(S_N) < 1$, while in our approach the condition for the stability is not directly related to the joint spectral radius of S_N.*

Proof: Using the same technique as in the proof of Theorem 6.1, replacing S_1 by S_N and remarking that ρ of hypothesis (44) is smaller than $m^{1/p}$, we immediately get:

$$\|v - \tilde{v}\|_{L^p} \lesssim \|v^0 - \tilde{v}^0\|_{\ell^p(\mathbb{Z}^d)} + \sum_{l \geq 0} m^{-l/p} \|d^l - \tilde{d}^l\|_{\ell^p(\mathbb{Z}^d)}, .$$

from which we deduce that:

$$\begin{aligned}\|v - \tilde{v}\|_{L^p(\mathbb{R}^d)} &\leq \|v^0 - \tilde{v}^0\|_{\ell^p(\mathbb{Z}^d)} + \|(m^{(s/d-1/p)j} \|d^j - \tilde{d}^j\|_{\ell^p(\mathbb{Z}^d)})_{j \geq 0}\|_{\ell^q(\mathbb{Z}^d)} \|(m^{-sj/d})_{j \geq 0}\|_{\ell^{q'}(\mathbb{Z}^d)} \\ &\lesssim \|v^0 - \tilde{v}^0\|_{\ell^p(\mathbb{Z}^d)} + \|(m^{(s/d-1/p)j} \|d^j - \tilde{d}^j\|_{\ell^p(\mathbb{Z}^d)})_{j \geq 0}\|_{\ell^q(\mathbb{Z}^d)},\end{aligned}$$

It remains to estimate the semi-norm

$$|w|_{B^s_{p,q}(\mathbb{R}^d)} := \|(m^{js/d} \omega_N(w, m^{-j/d})_{L^p})_{j \geq 0}\|_{\ell^q(\mathbb{Z}^d)},$$

for $w := v - \tilde{v}$. For every $j \geq 0$, denoting $w_j = v_j - \tilde{v}_j$, we have

$$\omega_N(w, m^{-j/d})_{L^p} \leq \omega_N(w - w_j, m^{-j/d})_{L^p} + \omega_N(w_j, m^{-j/d})_{L^p}. \qquad (45)$$

For the first term, using successively Lemma 6.1, hypothesis (44), and then making the same kind of computation as in (32)one has

$$\begin{aligned}
\omega_N(w - w_j, m^{-j/d})_{L^p} &\lesssim \sum_{l\geq j} \|w_{l+1} - w_l\|_{L^p(\mathbb{R}^d)} \\
&\lesssim \sum_{l\geq j} m^{-l/p} \left(\|\Delta^N(v^l - \tilde{v}^l)\|_{(\ell^p(\mathbb{Z}^d))^{q_N}} + \|d^l - \tilde{d}^l\|_{\ell^p(\mathbb{Z}^d)} \right) \\
&\lesssim \sum_{l\geq j} m^{-l/p} \left(\rho^l \|v^0 - \tilde{v}^0\|_{\ell^p(\mathbb{Z}^d)} + \sum_{k=0}^{l} \rho^{l-k} \|d^k - \tilde{d}^k\|_{\ell^p(\mathbb{Z}^d)} \right).
\end{aligned}$$

Since $\rho < m^{1/p}$, then for the first term we have

$$\sum_{l\geq j} m^{-l/p} \rho^l \|v^0 - \tilde{v}^0\|_{\ell^p(\mathbb{Z}^d)} \sim m^{-j/p} \rho^j \|v^0 - \tilde{v}^0\|_{\ell^p(\mathbb{Z}^d)}.$$

For the second term, we get

$$\begin{aligned}
&\sum_{l\geq j} m^{-l/p} \sum_{k=0}^{l} \rho^{l-k} \|d^k - \tilde{d}^k\|_{\ell^p(\mathbb{Z}^d)} = \\
&\quad = m^{-j/p} \sum_{k=0}^{j} \rho^{j-k} \|d^k - \tilde{d}^k\|_{\ell^p(\mathbb{Z}^d)} + \sum_{l>j} m^{-l/p} \sum_{k=0}^{l} \rho^{l-k} \|d^k - \tilde{d}^k\|_{\ell^p(\mathbb{Z}^d)} \\
&\quad \lesssim m^{-j/p} \sum_{k=0}^{j} \rho^{j-k} \|d^k - \tilde{d}^k\|_{\ell^p(\mathbb{Z}^d)} + \sum_{k\geq 0} \sum_{l\geq \max(k,j+1)} m^{-l/p} \rho^{(l-k)} \|d^k - \tilde{d}^k\|_{\ell^p(\mathbb{Z}^d)} \\
&\quad \lesssim m^{-j/p} \sum_{k=0}^{j} \rho^{j-k} \|d^k - \tilde{d}^k\|_{\ell^p(\mathbb{Z}^d)} + \sum_{k>j} m^{-k/p} \|d^k - \tilde{d}^k\|_{\ell^p(\mathbb{Z}^d)}.
\end{aligned}$$

The second term in (45) is evaluated as follows:

$$\begin{aligned}
\omega_N(v_j - \tilde{v}_j, m^{-j/d})_{L^p} &\lesssim \|v_j - \tilde{v}_j\|_{L^p(\mathbb{R}^d)} \\
&\lesssim m^{-j/p} \rho^j \|v^0 - \tilde{v}^0\|_{\ell^p(\mathbb{Z}^d)} + m^{-j/p} \sum_{l=0}^{j} \rho^{j-l} \|d^l - \tilde{d}^l\|_{\ell^p(\mathbb{Z}^d)}.
\end{aligned}$$

The second term on the right hand side of (45), can be evaluated the same way. We have thus reduced the estimate of $|w|_{B^s_{p,q}(\mathbb{R}^d)}$, to the estimates of the discrete norms $\|(m^{js/d}a_j)_{j\geq 0}\|_{\ell^q(\mathbb{Z}^d)}$, and $\|(m^{js/d}b_j)_{j\geq 0}\|_{\ell^q(\mathbb{Z}^d)}$, where the sequences are defined as follows:

$$\begin{aligned}
a_j &:= \rho^j m^{-j/p}\|v^0 - \tilde{v}^0\|_{\ell^p(\mathbb{Z}^d)}, \\
b_j &:= m^{-j/p}\sum_{l=0}^{j}\rho^{j-l}\|d^l - \tilde{d}^l\|_{\ell^p(\mathbb{Z}^d)}, \\
c_j &:= \sum_{l>j} m^{-l/p}\|d^l - \tilde{d}^l\|_{\ell^p(\mathbb{Z}^d)}.
\end{aligned}$$

Note that this quantities are identical to that obtained in the convergence theorem replacing v^l by $v^l - \tilde{v}^l$ and d^l by $d^l - \tilde{d}^l$, so that the end of the proof is identical. □

7. Conclusion

In this paper, we have presented a new kind of nonlinear and non-separable multi-scale representations based on the use of non-diagonal dilation matrices. We have shown that the non-linear scheme proposed by Harten naturally extends in that context and we have shown convergence and stability in L^p and Besov spaces. The key idea is to use the characterization of Besov spaces by means of mixed finite differences and then to study the joint spectral radii of these difference operators. Because, these operators involve all potential mixed finite differences their study cannot be reduced to that of one-dimensional difference operators. Future work should involve the study of multi-scale representations which are associated with more general prediction operators which do not necessarily exactly reproduce polynomials.

[1] Amat, S., Arandiga, F.,Cohen,A. and Donat,R., Tensor product mul-

tiresolution analysis with error control for compact image representations, Signal Processing, 82, 587-608, (2002).

[2] Arandiga, F.,Cohen,A., Donat,R., Dyn, N. and Mateï,B., Approximation of piecewise smooth images by edge-adapted techniques, ACHA, 24, 225-250, (2008).

[3] E.J. Candès and D. L. Donoho New tight frames of curvelets and optimal representations of objects with piecewise-C2 singularities, Comm. Pure Appl. Math., 57, 219-266, (2002).

[4] Chan, T.F. and Zhou, H.M., ENO-Wavelet Transforms for Piecewise Smooth Functions, SIAM Journal on Numerical Analysis, 40 , Issue 4 , 1369 - 1404, (2002).

[5] Chappelier V. and Guillemot C., Oriented Wavelet for Image Compression and Denoising, IEEE Transactions on Image Processing, 15, 2892-2903, (2006).

[6] Cohen,A., Numerical Analysis of Wavelet Method, Elsevier Publisher, (2003).

[7] Do M.N. and Vetterli M., The contourlet transform: an efficient directional multiresolution image representation, IEEE Transactions Image on Processing, 14, no. 12, 2091-2106,(2005).

[8] Goodman,T.N.T., Michelli, C.A., and Ward,J.D., Spectral Radius Formulas for Subdivision Operators, Recent Advances in Wavelet Analysis, Schumaker, L.L.and Webb,G., eds., Academic Press, San Diego, 335-360, (1994).

[9] Harten, A., Discrete Multiresolution Analysis and Generalized Wavelets, Journal of Applied Numerical Mathematics, 12, 153-193 (1993).

[10] S. Harizanov and P. Ostwald, Stability of Nonlinear Subdivision Multiscale Transform, Constr. Approx., 31, 359-393, (2010).

[11] Harten, A., Enquist, B., Osher, S. and Chakravarthy, S., Uniformly High Order Accurate Essentially Non-oscillatory Schemes III, Journal of Comput. Phys., 71, 231-303, (1987).

[12] Jia, R-Q., Characterization of Smoothness of Multivariate Refinable Fucntions in Sobolev Spaces, Transactions of the American Mathematical Society, 49 , 4089-4112, (1999).

[13] LePennec, E., and Mallat, S., Sparse Geometrical Image Approximation with Bandelets, IEEE Transaction on Image Processing, 14,423-438, (2004).

[14] Malassiotis S., Strintzis M.G., Optimal Biorthogonal Wavelet Decomposition of Wire-Frame Meshes Using Box Splines and its Application to the Hierarchical Coding of 3D Surfaces, IEEE Transactions on Image Processing, 8, 41-57, (1999).

[15] Mallat, S., A Wavelet Tour of Signal Processing, 3rd ed., Third Edition: The Sparse Way, Academic Press, 3rd edition, (2008).

[16] Mateï, B., Méthodes multi-échelles non-linéaires–Applications au traitement d'images, Ph.D. Thesis, Université Pierre et Marie Curie (Paris VI), (2002).

[17] Mateï, B., Smoothness Characterization and Stability in Nonlinear Multiscale Framework: Theoretical Results, Asymptotic Analysis, 46, 277-309, (2005).

[18] Mateï,B., Meignen,S., Zakharova,A., Smoothness of Nonlinear and Nonseparable Subdivision Schemes, submitted, hal-00452897.

[19] Dyn,N., Oswald, P., Univariate Subdivision and Multi-Scale Transforms: the Nonlinear Case, in Multiscale, Nonlinear and Adaptive Approximation, Springer, 203-247, (2009).

[20] Storozhenko, E., Oswald, P., Jackson's Theorem in the spaces $L^p(\mathbb{R}^k), 0 < p < 1$, Sibirsk. Math. Zh., 19, 888-901; English transl. Siberian Math. J., 19 (1978).

[21] Shapiro, J. M., Embedded Image Coding Using Zerotrees of Wavelet Coefficients, IEEE Transactions on Signal Processing, 41, No. 12, p. 3445-3462, 1993.